



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/803,478	03/17/2004	Yih-Kuen Tsay	TP0010-NTU-003-US	5502
64686	7590	04/18/2007		
JOHN CHEN ROOM 303, 3F, NO. 25, SEC. 1, CHANGAN E. ROAD TAIPEI (R.O.C.), 10441 TAIWAN			EXAMINER RAYYAN, SUSAN F	
			ART UNIT 2167	PAPER NUMBER
SHORTENED STATUTORY PERIOD OF RESPONSE		MAIL DATE	DELIVERY MODE	
3 MONTHS		04/18/2007	PAPER	

**Please find below and/or attached an Office communication concerning this application or proceeding.**

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

# Office Action Summary

Application No.

10/803,478

Applicant(s)

TSAY ET AL.

Examiner

Susan F. Rayyan

Art Unit

2167

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1) ☒ Responsive to communication(s) filed on 24 January 2007.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

- 4) ☒ Claim(s) 1-20 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-20 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

## Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

## Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

## Attachment(s)

- 1) ☐ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: \_\_\_\_\_

## DETAILED ACTION

1. Claims 1-20 are pending.

### ***Claim Rejections - 35 USC § 103***

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

**Claims 1, 3-16 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Number 6,704,728 issued to Jane Wen Chang et al ("Chang") in view of US Patent Publication 2004/0039734 issued to Douglas Russell Judd et al ("Judd").**

**As per independent claim 1 Chang teaches:**

A sequence based indexing and retrieval method for text documents (Abstract), comprising the steps of:

(a) generating a query token sequence, having at least a query token, from a query submitted by a user at column 2, lines 56-67 as a query is a word, multiple words or sentence fragment. The query is normalized and expanded (column 9, lines 60-62) into the claimed "query token sequence". Specification indicated a token as a word or character (page 1, line 14);

Art Unit: 2167

(b) generating at least a representative token sequence, having at least a document token, from each of said text documents that contain at least one token of said query token sequence at column 1, lines 39-48 as documents containing words are parsed and grammar rules applied to generate an IR index containing "document tokens". Each document has associated with it the document tokens. A query search would produce a "representative token sequence" of a document when a query token matches a document token in the index.

Chang does not explicitly teach measuring a similarity between each of said representative token sequences and said query token sequence by: (c.1) determining a token appearance score by measuring a token appearance of said representative token sequence with respect to said query token sequence; (c.2) determining a token order score by measuring a token order of said representative token sequence with respect to said query token sequence; and (c.3) determining a token consecutiveness score by measuring a token consecutiveness of said representative token sequence with respect to said query token sequence; and (d) retrieving said text documents in responsive to said similarity of said representative token sequence with respect to said query token sequence with a ranking order in accordance with said token appearance score, said token order score, and said token consecutiveness score, provided that for a document with two representative token sequences, its similarity is determined by the representative token sequence with a higher score.

Art Unit: 2167

Judd does teach these measuring a similarity between each of said representative token sequences and said query token sequence by:

(c.1) determining a token appearance score by measuring a token appearance of said representative token sequence with respect to said query token sequence (paragraphs 56,58,68, as matching search terms);

(c.2) determining a token order score by measuring a token order of said representative token sequence with respect to said query token sequence (paragraph 26, lines 7-9 and paragraph 67 and desired search terms in same order are generally ranked higher);

(c.3) determining a token consecutiveness score by measuring a token consecutiveness of said representative token sequence with respect to said query token sequence (paragraph 57 and paragraph 64, as distance between matching search terms);

(d) retrieving said text documents in responsive to said similarity of said representative token sequence with respect to said query token sequence with a ranking order in accordance with said token appearance score, said token order score, and said token consecutiveness score, provided that for a document with two representative token sequences, its similarity is determined by the representative token sequence with a higher score ( paragraph 10) to help users locate a desired document. It would have been obvious to one of ordinary skill in the art at the time of the invention to modify Chang with determining a token order score, determining a token appearance score, determining a token

Art Unit: 2167

consecutiveness score to help users locate a desired document at paragraph 8.

**As per claim 3**, same as claim arguments above and Judd teaches:

wherein said weight of said query token in said query token sequence is measured by determining a token frequency of said query token in said text documents (paragraph 70).

**As per claim 4,5**, same as claim arguments above and Judd teaches:

determining a length of the longest common subsequence of said representative token sequence and said query token sequence, determining a length of said representative token sequence, determining a length of said query token sequence, and outputting said token order score of said token order by calculating a fraction of said length of said longest common subsequence divided by an average sum of said length of said representative token sequence and said length of said query token sequence (paragraph 70).

**As per claim 6,7,8** same as claim arguments above and Judd teaches:

determining a relative distance between a positional differentiation of each adjacent document tokens and a positional differentiation of said adjacent document tokens in said query token sequence, and (c.3.2) outputting said token consecutiveness score of said token consecutiveness by calculating a fraction of a sum of the inverses of said relative distances divided by the number of pairs of adjacent tokens, which equals the length of said representative token sequence less one(paragraph 56-paragraph 64).

**As per claim 9**, same as claim arguments above and Judd teaches:

Art Unit: 2167

wherein said similarity of said representative token sequence is calculated with respect to said query token sequence by summing said token appearance score, said token order score, and said token consecutiveness score, wherein said ranking order of said text documents is determined by a weighted sum of said token appearance score, said token order score, and said token consecutiveness score of each of said representative token sequences of said text documents (paragraphs 56, 58, 68, as matching search terms, paragraph 26, lines 7-9, paragraph 67 as desired search terms in same order are generally ranked higher, paragraph 57 and paragraph 64, as distance between matching search terms).

**As per claim 10, 11** same as claim arguments above and Judd teaches:

further comprising a step of selecting at least a candidate document from said text documents, wherein one of said text documents is selected to be said candidate document when said text document contains at least one token of said query token sequence (paragraph 55).

**As per claim 12, 13** same as claim arguments above and Judd teaches:

further comprising a step of consulting an index of said text documents to establish said candidate document, wherein tokens that also appear in the query token sequence are collected to form a document token sequence for each document and the two longest segments of said document token sequence are selected as representative token sequences wherein the positional differentiation of each adjacent document tokens is no larger than a predetermined positioning value while said corresponding text document is selected as the said candidate

Art Unit: 2167

document (paragraph 50, indexed documents and paragraph 57, search term proximity).

**As per claim 14, 15,16** same as claim arguments above and Judd teaches: further comprising a step of retaining said candidate document to be used for measuring said similarity with respect to said query token sequence, wherein the said candidate document is retained when said candidate document contains a token that has a weight no less than a predetermined fraction of the total weight of query tokens (paragraph 55).

**Claim2 is rejected under 35 U.S.C. 103(a) as being unpatentable over Chang in view of Judd as applied to claim 1 above, and further in view of US Patent Publication 2003/0028520 issued to Shamim A. Alpha et al ("Alpha").**

**As per claim 2**, same as claim arguments above and Chang in view of Judd teach , calculating a sum of the weights of the query tokens that appear in said representative token sequence and outputting said token appearance score of said token appearance by calculating a fraction of said sum divided by the total weight of all query tokens (paragraph 10). Chang in view of Judd do not explicitly teach consulting an index of said text documents to determine the weight of each token in said query token sequence. Alpha does teach this limitation to improve query response time and improve search results. It would have been obvious to a person of ordinary skill in the art at the time of the invention was made to modify Chang in view of Judd with consulting an index of said text documents to determine the weight of each token in said query token sequence to improve



Art Unit: 2167

query response time and improve search results as described by Alpha

(Abstract).

**Claims 17-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chang in view of Judd as applied to claims 1-3,5,8,9,11,13,16 above, and further in view of US 2002/0022953 issued to Phillip Andre Bertolus et al ("Bretolus").**

As per claim 17,18,19,20 same as claim arguments above and Chang in view of Judd do not explicitly teach wherein said text document contains Chinese characters, English words, numbers, punctuations, and symbols as said document tokens. Bertolus does teach this limitation at paragraph 38 to retrieve and index information on a network. It would have been obvious to a person of ordinary skill in the art at the time of the invention was made to modify Chang and Judd with text document contains Chinese characters, English words, numbers, punctuations, and symbols as said document tokens to retrieve and index information on a network as described by Bertolus (paragraph 3).

#### Response to Arguments

3. Applicant's arguments filed January 24, 2007 have been fully considered but they are not persuasive.

Applicant argues prior art of record does not teach "generating a query token". Examiner finds Chang teaches this limitation at column 2, lines 56-67 as a query is a word, multiple words or sentence fragment. The query is normalized and expanded (column 9, lines 60-62) into the claimed "query token sequence". Specification indicated a token as a word or character (page 1, line 14).

Art Unit: 2167

Applicant argues prior art of record does not teach "generating at least a representative token sequence having at least a document token, from each of said query token sequence. Chang does teach this limitation at column 1, lines 39-48 as documents containing words are parsed and grammar rules applied to generate an IR index containing "document tokens". Each document has associated with it the document tokens. A query search would produce a "representative token sequence" of a document when a query token matches a document token in the index.

Applicant argues prior art of record does not teach "token appearance score". Judd teaches this limitation at paragraphs 56, 58, 68, as relevance ranking uses heuristics such as matching search terms (tokens) to judge match quality.

Applicant argues prior art of record does not teach "determining a token order score by measuring a token order of said representative token sequence with respect to said query token sequence". Judd teaches this limitation at paragraph 26, lines 7-9 and paragraph 67 and desired search terms (tokens) in same order are generally ranked higher.

Applicant argues prior art of record does not teach "a token consecutiveness score". Judd teaches this limitation at paragraph 57 and paragraph 64, as documents having matching search terms (tokens) in close proximity help determine match quality.

***Conclusion***

4. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than **SIX MONTHS** from the mailing date of this final action.

Art Unit: 2167

### Contact Information


5. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Susan Rayyan whose telephone number is (571) 272-1675. The examiner can normally be reached M-F: 8am - 4:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Cottingham can be reached on (571) 272-7079. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

  
Susan Rayyan

April 11, 2007

  
JOHN COTTINGHAM  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2100